

**A Platform for Distributed Computational Biology****Maewal, Akhilesh\*, Bock, Joel R.****Intelligent Automation Corporation, Poway, CA, USA**

Bioinformatics algorithms and databases have become increasingly specialized in content and idiosyncratic in format in direct proportion to their rate of appearance on the World Wide Web. This situation creates “barriers to entry” to scientific researchers with neither the requisite background nor inclination to create and maintain computer programs to access these important bioinformatics resources—such barriers directly obstruct the advancement of scientific knowledge. Ironically, many of these public domain resources have been developed using public funds, yet they remain essentially unusable by the majority of potential users.

A partial solution to this dilemma is to use a “web services” approach to make online databases resource programmatically available via common interfaces [1]. This has the potential to solve the problem of communication between a single researcher and a data resource. An open issue is the need for an information architecture that facilitates open collaboration and the shared use of bioinformatics visualization, analysis and data mining tools and algorithms *between individuals*. For example, most bioinformatics researchers are able to perform BLAST searches using web browser-based access to remote services. However, bioinformatics entails much more than homology search. More complicated computations, perhaps involving scripts to assemble several programs in an analysis “pipeline”, cannot be carried out without specific expertise in the programming language(s) used to implement each particular code, for example. We submit that the underlying biological concept embodied within an algorithm is fundamentally and intrinsically valuable—not the programming language used in its implementation. All bioinformatics application software should be universally accessible to support scientific research.

We are developing an extendible and comprehensive system for collaborative biological computations. The system is based upon Sun’s JavaSpaces<sup>®</sup> technology implemented on the Jini platform. JavaSpaces technology is a simple unified mechanism for dynamic communication, coordination, and sharing of objects that encapsulate biological data, models, and algorithms, between Java technology-based network resources like clients and servers. In our distributed application, JavaSpaces technology acts as a virtual space between providers and requesters of data or objects. This allows participants in a distributed solution to exchange tasks, requests and information in the form of Java technology-based objects. JavaSpaces technology provides developers with the ability to create and store objects with persistence, which allows for process integrity. The browser based clients allows the users to access the system transparently to (a) access data from heterogeneous sources, (b) develop models, and (c) perform computations by using distributed resources.

**Reference**

L. Stein, “Creating a bioinformatics nation”, *Nature* 417:119-120, 2002.